CHROMSYMP. 1924

Quantitative correlations between solute molecular structure and solubility in supercritical fluids

JERRY W. KING* and JOHN P. FRIEDRICH

Northern Regional Research Center, Agricultural Research Service, United States Department of Agriculture, 1815 North University Street, Peoria, IL 61604 (U.S.A.)

ABSTRACT

There is a current need in the fields of supercritical fluid extraction and supercritical fluid chromatography (SFC) for rapid and simple methods of estimating a solute's solubility level in the dense gaseous solvent. Utilizing the solubility parameter theory developed by Giddings, we have developed a method which permits the quantitative estimation of solute solubility levels in dense and liquefied gas media over a range of pressure and temperatures. The described method incorporates the ratio of the solubility parameter of the extraction gas to that of the dissolved solute, thereby permitting correlations to be made for a number of solute–gas combinations. Techniques are also presented that permit solute solubilities to be estimated from a knowledge of the solute's molecular structure. The above methods have been applied to such applications as the extraction of organic solutes from aqueous media with liquefied carbon dioxide and supercritical fluid chromatography.

INTRODUCTION

The modern era of supercritical fluid extraction (SFE) has spawned a wide number of diverse applications in the areas of natural product extraction and food processing. As in many fields of technological endeavor, theoretical developments in supercritical fluid technology have trailed experimental and application studies. This trend is most unfortunate since there is a critical need for basic thermodynamic and kinetic models upon which to optimize extraction and separation conditions. Current theoretical developments in this field have included the application of statistical mechanical models[1,2], equation of state methods[3,4] and a diverse array of solution thermodynamic concepts[5-7] to explain and correlate phase equilibria and solubility phenomena in supercritical fluids. In general, the above approaches require extensive physicochemical data to characterize solute–solvent (gas) interactions and they have been applied mostly to a limited number of structurally-simple solutes dissolved in supercritical carbon dioxide. Such theoretical methods currently find limited use in practical applixcations, in part due to the molecular complexity of the solutes that are

being extracted and the lack of physical property data on these moieties.

Over the past seven years. we have applied solubility parameter theory[8]. interaction parameters[9] and corresponding states methods to explain the salient features of SFE. One advantage of the above approach is the availability of required data for both the solvent gas and extracted solutes that are required by the regular solution or corresponding states theory. Utilizing critical property data and solute–solvent solubility parameters. one can estimate the required gas pressures (or density) to affect maximum solute solubility in the dense gas phase. Prediction of this solubility maximum can be determined by[10]

$$\chi = \bar{V}_1 (\delta_1 - \delta_2)^2 / RT + \chi_s \qquad (1)$$

where χ is the total interaction parameter. χ_s is the entropic interaction parameter. δ_1 is the solubility parameter of the gas as a $f(T, P)$, δ_2 is the solubility parameter of the solute as a $f(T, P)$, and $\bar{V}_1$ is the molar volume of the gas as a $f(T, P)$. R is the molar gas constant, T is the absolute temperature and P is the pressure.

Similarly, miscibility of the solute in the supercritical fluid can be predicted to occur at a discrete pressure, which can be ascertained by solving for the intersection of χ and χ_c as a function of pressure, where the Flory critical interaction parameter. χ_c, is given by

$$\chi_c = (1 + x^{\frac{1}{2}})^2 / 2x \qquad (2)$$

and $x = \bar{V}_2 / \bar{V}_1$, where $\bar{V}_2$ is the molar volume of the solute.

The above equations have been tested in a number of cases involving supercritical fluid equilibria and have been found to predict results that agree with experimental data. provided that the temperature and pressure dependence of solubility parameters and molar volumes are utilized in the calculations. Pertinent applications have included vegetable oil extraction[11] polymer solution thermodynamics[12] and supercritical fluid chromatography (SFC)[10,13].

HISTORICAL BACKGROUND

Correlation of a solute's molecular structure with its solubility in a supercritical fluid has been qualitatively attempted by several investigators. Stahl et al.[14]. utilizing the results from his extensive micro-extraction studies. formulated several extraction rules based on changes in a solute's molecular structure to qualitatively predict the extent of a solute's dissolution in supercritical carbon dioxide. Qualitative trends in solute solubilities have also been summarized by Hyatt[15] and by Dandge et al.[16] who supplemented the extensive liquid carbon dioxide data of Francis[17] with additional data from their experimental studies. A concise summary of solubility in liquid carbon dioxide has also been reported by Sims[18]. In a recent extensive review Rizvi et al.[19] have discussed the current state of knowledge regarding solute solubilities in supercritical fluid media.

There is currently a need for methods that allow the rapid quantitative assessment of a solute's solubility in a supercritical fluid. Such information is critical for rendering decisions regarding the feasibility of applying supercritical fluid

extraction to an array of engineering and analytical problems. As noted by McHugh and Krukonis[20], there have been several bizarre attempts to apply SFE to the extraction of compounds exhibiting limited solubility in supercritical fluid media. Such questionable applications could be avoided if correlations existed between a solute's molecular structure and its recorded solubility in a particular supercritical fluid. The development of simple predictive methods to address the above problem is difficult due to the limited solubility data bases available and the diversity of units in which solute solubility data are frequently reported.

THE REDUCED SOLUBILITY PARAMETER CONCEPT

In this study, we have utilized the concept of the reduced solubility parameter, Δ, to correlate solubilities and their distribution coefficients in supercritical and near-critical fluids under a variety of extraction conditions. The reduced solubility parameter is defined as[21]

$$\Delta = \delta_1/\delta_2 \tag{3}$$

where δ_1 can be calculated by methods described below and δ_2 is determined from standard sources[22]. It should be noted that this definition of "reduced solubility parameter" differs from that utilized by Prausnitz[23] and Allada[24] who defined the term as the solubility parameter of the gas divided by the square root of the gas critical pressure. The ratio of the solubility parameter of the solvent (gas) to that of the solute acutely reflects the strength of the solute–solvent interactions under a variety of pressure and temperature conditions and was first suggested by Giddings et al.[21] in 1969. As an aid in our correlations we have employed the group contribution method of Fedors[25] to compute the solubility parameters for many structurally complex solutes such as carotenoids, alkaloids, pesticides, sterols and antibiotics. By incorporating this latter concept into our method, we have made δ_2 a function of molecular structure and therefore Δ sensitive to changes in solute structure relative to the solvent gas.

To describe the solvent power of the dense gas, we have used the equation proposed by Giddings et al.[26] in which the solubility parameter of the gas is given by

$$\delta_1 = 1.25 \, P_c^{\ddagger}(\rho_{r,\,SF})/(\rho_{r,\,liquid}) \tag{4}$$

where P_c is the fluid critical pressure, $\rho_{r,\,SF}$ the reduced density of the critical fluid, and $\rho_{r,\,liquid}$ is the reduced density of the fluid in the liquid state. This equation has been shown to reflect the variance in solvent power of the gas as a function of pressure and temperature and is dependent on the chemical nature of the extraction fluid through P_c. Reduced densities for a variety of fluids can be determined by using gas densities and the appropriate critical constant, or in the absence of actual gas density data, through corresponding states relationships. Reduced densities of fluids in their liquid state are computed at infinite pressure and the proper reduced temperature from corresponding states tables.

RESULTS AND DISCUSSION

Solute solubilities in supercritical fluids depend not only on a variance in the molecular structure, between solutes, but also on the pressure and temperature at which the extraction is performed. Such factors can make a solute's solubility in the dense gas phase vary over many orders of magnitude. However, in general, solubilities recorded in supercritical fluids are less than those recorded in liquid solvents. An example of this solubility trend is shown in Fig. 1 where weight fraction solubilities for various solutes dissolved in supercritical fluid ethylene are plotted against the reduced solubility parameters Δ. The solubility data were obtained from the "classical Zosel" patent issued to SGK[27] and encompass a variety of solutes of varying polarity the solubility parameters of which are known. The reported solubility data in ethylene for these compounds encompassed a wide range of reduced pressures (1.19–3.96) and essentially isothermal conditions (20–25°C). The wide variation in reported solubilities is conveniently handled by using a logarithmic concentration scale for the ordinate while the variation in solute solubility parameters (7.77–14.4 $cal^{1/2}/cm^{3/2}$) and gas solubility parameters (3.84–6.4 $cal^{1/2}/cm^{3/2}$) is encompassed in the reduced solubility parameter scale on the abscissa. Note that as the ratio of the solubility parameter of the gas to that of solute increases, there is a substantial increase in the recorded solubility up to a reduced solubility parameter of 0.5. To a large extent, the recorded trend in the solubility data is reflective of the changing chemical nature of the solute molecule with maximum solubilities being recorded as the reduced solubility parameter tends to unity. The data shown in Fig. 1 appear to obey a functional form that can be conveniently fitted by a polynomial equation over the entire range of experimental data.

As indicated in Fig. 1, solute solubilities recorded at reduced solubility parameter values of less than 0.5 are extremely sensitive to changes in solute molecular structure. The results presented in Fig. 2 for supercritical carbon dioxide–solute systems further corroborate this trend, even among polar solutes of varying molecular
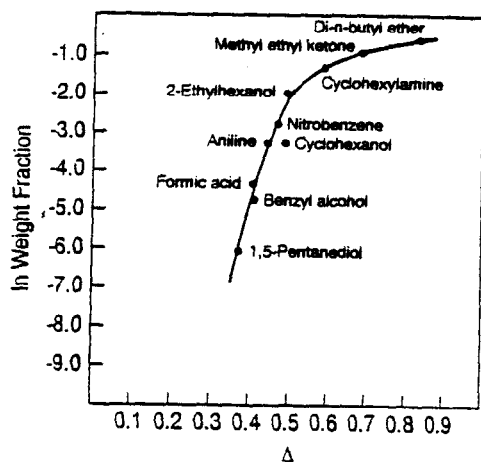
Fig. 1. Natural logarithm of solute weight fraction solubility in ethylene vs. reduced solubility parameter (based on experimental solubility data by Zosel[27]).
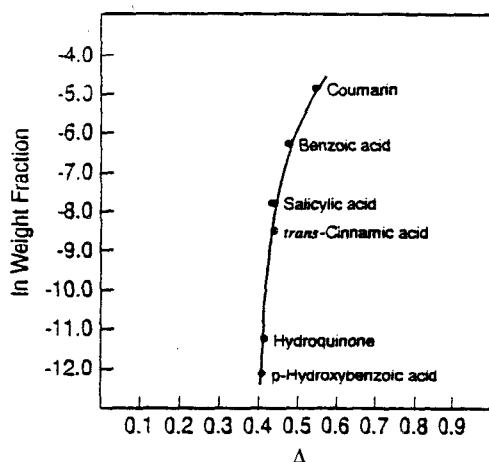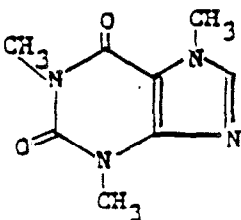
Fig. 2. Natural logarithm of solute weight fraction solubility in carbon dioxide (100 atm. 40°C) *vs.* reduced solubility parameter.

structure. The recorded weight fractions for the compounds (taken at 100 atm and 40°C), in Fig. 2 were interpolated from the studies of Schilz[28], while the solute solubility parameters were computed by the Fedors method. Fig. 2 shows the relationship between the natural logarithm of the weight fraction solubility and the corresponding reduced solubility parameter Δ for the designated solutes in carbon dioxide at 100 atm and 40°C. The data presented range in magnitude from a weight fraction of 7.7 · 10^{-3} for coumarin (δ_2 = 10.5 cal$^{1/2}$/cm$^{3/2}$) to a value of 5.5 · 10^{-6} for polar *p*-hydroxybenzoic acid (δ_2 = 13.94 cal$^{1/2}$/cm$^{3/2}$). A similar correlation has also been found to hold for phenolic solutes dissolved in supercritical carbon dioxide in which the solubility parameters of the solutes ranged from 13.96 to 21.4 cal$^{1/2}$/cm$^{3/2}$. Such correlations are valuable for they permit the rapid estimation of solute solubilities under similar extraction conditions when knowledge of the solute's solubility parameter and molecular structure are available.

TABLE I

CALCULATION OF THE SOLUBILITY PARAMETER FOR CAFFEINE

	Group	$\Delta\varepsilon_i$ (cal/cm^3)	Δv_i (cm^3/mol)
	3 CH$_3$	3375	100.5
	2 C=	2060	−11.0
	1 CH=	1030	13.5
	3 N	3000	−27.0
	1 −N=	2800	5.0
	2 C=O	8300	21.6
	2 5–6-member rings	500	32.0
δ_2 = 12.96	2 conjugated double bonds	800	−4.4
		21 865	130.2

An example of the above approach using the group contribution method of Fedors is given in Table I. Here the solubility parameter for caffeine has been computed by summing the individual group contributions for the energy of vaporization $\Delta\varepsilon_i$ and molar volume Δv_i to the overall structure of the solute molecule. The solute's solubility parameter is calculated by taking the square root of the ratio of the sum of the energy contributions over the sum of respective group molar volumes[24]. $\delta_2 = (\Sigma\Delta\varepsilon_i/\Sigma\Delta v_i)^{1/2}$. For caffeine. which has been extracted industrially with supercritical carbon dioxide. the solubility parameter value is 12.96 $cal^{1/2}/cm^{3/2}$. Interpolating in Fig. 2 using a reduced solubility parameter of 0.44 yields a value of $4.0 \cdot 10^{-4}$ for the weight fraction solubility of caffeine under the same extraction conditions. This value compares favorably with the experimentally determined value of $3.3 \cdot 10^{-4}$ determined by Schilz[28].

Fedors' method is extremely valuable in ascertaining the cohesional energy density of complex molecules. With this method one does not even have to depend on the type of relationships noted above to reach seminal conclusions regarding the potential for extracting a compound under SFE conditions. Applications of this group contribution method to bioactive molecules. such as imipenum and efrotomycin which have been extracted by Larson and King[29] with supercritical carbon dioxide at 40°C and 5000 p.s.i.g. ($\delta_1 = 8.2\ cal^{1/2}/cm^{3/2}$), supports the correspondence between solute solubility parameter and solute solubility levels in supercritical fluids. For example. efrotomycin, a large molecule exhibiting antibiotic activity. has a solubility parameter of 12.1 $cal^{1/2}/cm^{3/2}$ and a weight fraction solubility in the dense gas phase of $3.0 \cdot 10^{-4}$. By comparison. imipenum. whose molar volume is 6.5 times smaller then efrotomycin's has a solubility parameter of 16.4 $cal^{1/2}/cm^{3/2}$ and exhibits no recorded solubility in carbon dioxide under the above-cited conditions. Good agreement is also recorded when one compares solubility parameters computed from Fedors' method with those obtained from other sources. As shown in Table II. the solubility parameters determined by Fedors' method for sterols agree well with the values given by Wong and Johnston[30].

Reduced solubility parameter correlations can provide a better understanding of the solubility trends in liquefied gases. Fig. 3 shows the correlation between Δ and the molar distribution coefficients. K (m). of various solutes partitioning between water and liquefied carbon dioxide held at 16°C ($\delta_1 = 7.15\ cal^{1/2}/cm^{3/2}$). The distribution coefficient data presented in Fig. 3 were taken from the studies of Schultz and Randall[31]. in which odoriferous alcohols and esters were recovered from aqueous

TABLE II

COMPARISON OF SOLUBILITY PARAMETERS CALCULATED BY TWO DIFFERENT METHODS FOR STEROLS

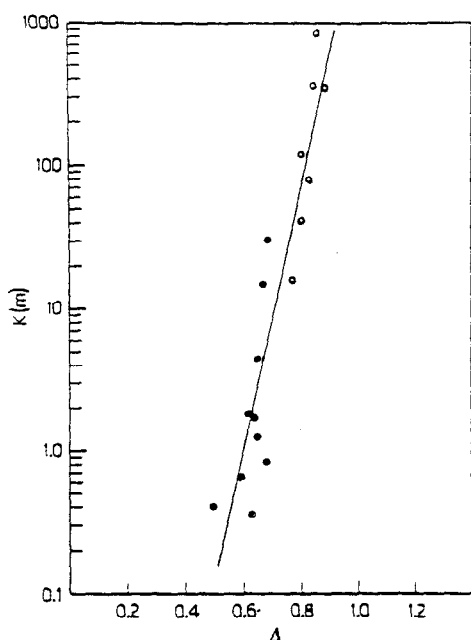| Sterol	Solubility parameter ($cal^{1/2}/cm^{3/2}$)	
	Wong and Johnston[30]	Fedors[25]
Cholesterol	9.2	9.6
Stigmasterol	8.8	9.5
Ergosterol	9.1	9.4

Fig. 3. Molar distribution coefficient *vs.* reduced solubility parameter for solutes partitioning between water and liquid carbon dioxide (16°C). Symbols: ● = alcohols; ○ = esters.

media by partitioning into liquid carbon dioxide. The semilogarithmic relationship between distribution coefficients and Δ is exemplified by the fit of a linear regression line to the data in Fig. 3. The value of such a correlation becomes apparent when one is considering the recovery of other candidate solutes from aqueous media. For example, the small distribution coefficients (0.002–0.010) found for polar solutes such as acetic acid, dimethylformamide and ethylene glycol are predicted by use of the reduced solubility parameter scale. Extraction of these solutes into liquefied carbon dioxide is somewhat impractical, in part due to the high solvent-to-feed ratio requirements. Nevertheless, their removal by critical fluid extraction has been proposed and cited in the literature[32,33].

The comprehensive solubility studies performed in liquefied carbon dioxide by Francis[17] over three decades ago have been cited by numerous investigators as a qualitative guideline for predicting solute solubilities in supercritical carbon dioxide. Such a data base, consisting of over 260 binary systems, is potentially correlatable by using the reduced solubility parameter concept. Unfortunately, much of the solubility data consisted of visual observations concerning total or complete immiscibility, determined in an era when appropriate quantitative analytical methods were limited. However, for certain cited systems, compositions were determined for both the liquefied carbon dioxide and second component layers. By using the weight fraction compositions for each of the respective phases, we have been able to define a distribution coefficient $K(w)$ as:

$$K(w) = \frac{\text{weight fraction of solute in the liquefied carbon dioxide phase}}{\text{weight fraction of solute in the second phase}} \qquad (5)$$

which is a measure of solute (second component) partition into the liquid carbon dioxide phase. Thirty-four of these solute distribution coefficients for an array of solutes partitioning into liquefied carbon dioxide have been plotted *versus* Δ on a semi-logarithmic graph in Fig. 4. Upon initial examination of Fig. 4. the data appear to yield no recognizable trend. However, classification of the individual systems by type of compound and their respective solubility parameters allows some conclusions to be drawn. The filled circle symbols represent very polar solutes. such as glycols. amides and amines, the solubility parameter values of which fall between 13 and 18 $cal^{1/2}/cm^{3/2}$ and which have a Δ of 0.2–0.5. Distribution coefficients for these compounds appear generally to lie below 0.01, consistent with their low solubility in liquefied carbon dioxide. Solutes of intermediate polarity (open circles). appear to have a Δ that falls between 0.5 and 0.7. Compounds in this group include substituted aromatic moieties, halogenated solutes and alcohols. Finally, solutes having solubility parameters less than 9.0 $cal^{1/2}/cm^{3/2}$ (half-filled circles) tend to exhibit distribution coefficients greater than 0.1. These include hydrocarbons, which partition partially into liquefied carbon dioxide or exhibit total miscibility with the liquefied gas. When the results of Fig. 4 are viewed in the above context, one has only to obtain a solute's solubility parameter to ascertain that solute's propensity for partitioning into the liquefied carbon dioxide phase.

Currently, the field of SFC is undergoing a renaissance due largely to improvements in instrumentation that can conviently handle highly compressed supercritical fluid eluents. Here, as in the above cases, a knowledge of the relative solubilities or partitioning behavior of the injected solutes in the column is critical to maximizing component resolution in the minimum amount of analysis time. The
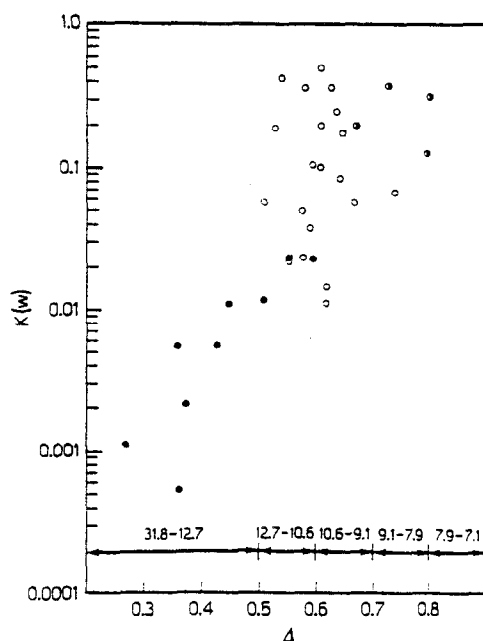

Fig. 4. Distribution coefficient *vs.* reduced solubility parameter for solutes partitioning into the liquid carbon dioxide phase (25°C).
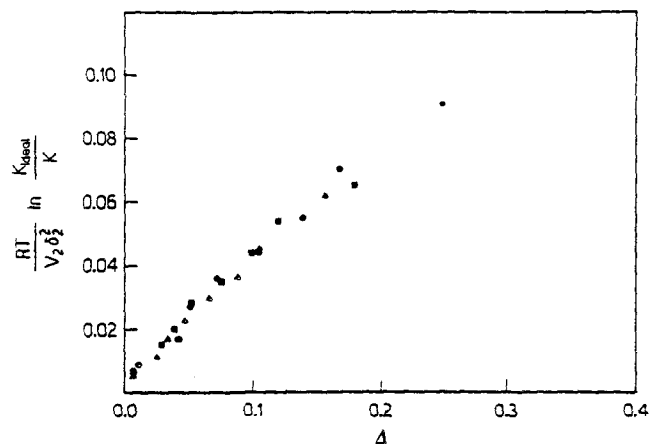
Fig. 5. $RT/\bar{V}_2\delta_2^2$ $\ln(K_{ideal}/K)$ vs. the reduced solubility parameter for solutes partitioning between supercritical carbon dioxide and squalane (40°C). Symbols: ● = benzene; ▲ = methanol; ■ = ethanol.

relative partitioning behavior of solutes with respect to each other and a low-pressure reference state (K_{ideal}) was formalized by Giddings et al.[21]

$$RT/\bar{V}_2\delta_2^2 \ln(K_{ideal}/K) = (2 - \Delta)\Delta \tag{6}$$

where K is the chromatographic partition coefficient. Eqn. 6 is expressed in reduced coordinates and predicts a general parabolic dependence of the solute's partition coefficient ratio on Δ. To illustrate the usefulness of the above concept, data interpolated from the studies of Sie et al.[34] have been used (Fig. 5) to construct a graph from Eqn. 6 for three solutes partitioning between squalane and carbon dioxide held at 40°C over the pressure range of 1–70 atm. By using the reduced solubility parameter as the ordinate of this figure, all of the data points for the solutes having significantly different solubility parameters can be made to conform to the same equation over the entire experimental pressure range. This conformity suggests that the partition coefficients could be predicted for additional solutes chromatographed under similar conditions if their reduced solubility parameters were computed. The above relationship has been shown to apply to other solute–column combinations, and its general applicability shows that pressure-induced changes in solute distribution coefficients are primarily independent of the chromatographic stationary phase. A similar method has been used by Ziger and Eckert[35] to describe solubility enhancement of solutes in supercritical fluids.

CONCLUSION

The cited examples show that it is possible to correlate a large body of supercritical fluid extraction data in a quantitative manner by relating the data to chnges in the molecular structure of the extracted solute. This can be accomplished readily by using the reduced solubility parameter concept as a measure of the solute–solvent interaction in the supercritical state. The advocated method is relatively

simple, requiring a minimum of experimental and physical data for the rapid estimation of solute solubilities in supercritical fluids. Such a concept is a valuable technical aid to the chromatographer faced with the need to make rapid decisions in utilizing SFE or SFC and for understanding the general solubility characteristics of these unique solvents.

REFERENCES

1 D. J. Vezzetti. *J. Chem. Phys.*. 77 (1982) 1512.
2 S. K. Kumar. R. C. Reid and V. W. Suter. in T. G. Squires and M. E. Paulaitis (Editors). *Supercritical Fluids: Chemical Engineering Principles and Applications*. American Chemical Society. Washington. DC. 1987. p. 88.
3 U. Deiters and G. M. Schneider. *Ber. Bunsen-Ges. Phys. Chem.*, 80 (1976) 276.
4 W. B. Whitting and J. M. Prausnitz. *Fluid Phase Equilib.*. 9 (1982) 119.
5 J. Chrastil. *J. Phys. Chem.*. 86 (1982) 3016.
6 R. T. Kurnik and R. C. Reid. *Fluid Phase Equilib.*. 8 (1982) 93.
7 G. A. Mansoori and J. F. Ely. *J. Chem. Phys.*. 82 (1985) 406.
8 J. W. King. *16th ACS Annual Great Lakes Regional Meeting, Normal. IL. June 9. 1982*. Abstract No. 245.
9 J. W. King and J. F. Friedrich. *20th ACS Annual Great Lakes Regional Meeting. Milwaukee. WI. June 2. 1986*, Abstract No. 170.
10 J. W. King. *J. Chromatogr. Sci.*. 17 (1989) 355.
11 J. W. King. *J. Am. Oil Chem. Soc.*. 60 (1983) 711.
12 J. W. King. *Polymeric Materials Sci. Eng. Preprints*. 51 (1984) 707.
13 J. W. King and J. F. Friedrich. *13th Annual Meeting of the Federation of Analytical Chemistry and Spectroscopy Societies. St. Louis. MO. Sept. 29. 1986*. Abstract No. 32.
14 E. Stahl. W. Schilz. E. Schultz and E. Willing. *Angew. Chem.. Int. Ed. Engl.*. 17 (1978) 731.
15 J. A. Hyatt. *J. Org. Chem.*. 49 (1984) 5097.
16 D. K. Dandge. J. P. Heller and K. V. Wilson. *Ind. Eng. Chem. Prod. Des. Dev.*. 24 (1985) 162.
17 A. W. Francis. *J. Phys. Chem.*. 58 (1954) 1099.
18 M. Sims. *Chem. Eng.*. 89 (1982) 50.
19 S. S. H. Rizvi. J. A. Daniels. A. L. Benado and J. A. Zollweg. *Food Technol. (Chicago)*. 40. No. 7 (1986) 57.
20 M. A. McHugh and V. J. Krukonis. *Supercritical Fluid Extraction*. Butterworths. Boston. MA. 1986.
21 J. C. Giddings. M. N. Myers and J. W. King. *J. Chromatogr. Sci.*. 7 (1969) 276.
22 A. F. M. Barton. *Handbook of Solubility and Other Cohesional Parameters*. CRC Press. Boca Raton. FL. 1983.
23 J. M. Prausnitz. in T. S. Storvick and S. I. Sandler (Editors). *Phase Equilibria and Fluid Properties in the Chemical Industry*. American Chemical Society. Washington. DC. 1977. p. 11.
24 S. R. Alada. *Ind. Eng. Chem. Process Des. Dev.*. 23 (1984) 344.
25 R. F. Fedors. *Polym. Eng. Sci.*. 14 (1974) 147.
26 J. C. Giddings. M. N. Myers. L. McLaren and R. A. Keller. *Science (Washington. D.C.)*. 162 (1968) 67.
27 K. Zosel. *U.S. Pat.*. 3 969 196 (1976).
28 W. Schilz. *Ph.D. Thesis*. University of Saarland. Saarbrücken. 1978.
29 K. A. Larson and M. L. King. *Biotechnol. Prog.*. 2 (1986) 73.
30 J. M. Wong and K. P. Johnston. *Biotechnol. Prog.*. 2 (1986) 29.
31 W. G. Schultz and J. M. Randall. *Food Technol. (Chicago)*. 24. No. 11 (1970) 94.
32 R. P. DeFilippi and J. M. Moses. in C. D. Scott (Editor). *Fourth Symposium on Biotechnology on Energy Production and Conservation*. Interscience. New York. 1982. p. 205.
33 E. J. Shimshick. *Chem. Tech.*. 13 (1983) 374.
34 S. T. Sie. W. van Beersum and G. W. A. Rijnders. *Sep. Sci.*. 1 (1966) 459.
35 D. E. Ziger and C. A. Eckert. *Ind. Eng. Chem. Process. Des. Dev.*. 22 (1983) 582.